

# 8



RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/745,506

DATE: 08/29/2002  
 TIME: 15:33:15

Input Set : N:\Crf3\RULE60\09745506.raw  
 Output Set: N:\CRF4\08292002\I745506.raw

## SEQUENCE LISTING

ENTERED

3 (1) GENERAL INFORMATION:  
 5 (i) APPLICANT: Lal, Preeti  
 6 Bandman, Olga  
 7 Hillman, Jennifer L.  
 8 Shah, Purvi  
 9 Corley, Neil C.  
 11 (ii) TITLE OF INVENTION: NEW HUMAN REGULATORY PROTEINS  
 13 (iii) NUMBER OF SEQUENCES: 74  
 15 (iv) CORRESPONDENCE ADDRESS:  
 16 (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 17 (B) STREET: 3174 Porter Drive  
 18 (C) CITY: Palo Alto  
 19 (D) STATE: CA  
 20 (E) COUNTRY: USA  
 21 (F) ZIP: 94304  
 23 (v) COMPUTER READABLE FORM:  
 24 (A) MEDIUM TYPE: Diskette  
 25 (B) COMPUTER: IBM Compatible  
 26 (C) OPERATING SYSTEM: DOS  
 27 (D) SOFTWARE: FastSEQ for Windows Version 2.0  
 29 (vi) CURRENT APPLICATION DATA:  
 C--> 30 (A) APPLICATION NUMBER: US/09/745,506  
 C--> 31 (B) FILING DATE: 21-Dec-2000  
 32 (C) CLASSIFICATION:  
 34 (vii) PRIOR APPLICATION DATA:  
 35 (A) APPLICATION NUMBER: US/08/870,870  
 36 (B) FILING DATE:  
 38 (viii) ATTORNEY/AGENT INFORMATION:  
 39 (A) NAME: Billings, Lucy J.  
 40 (B) REGISTRATION NUMBER: 36,749  
 41 (C) REFERENCE/DOCKET NUMBER: PF-0300 US  
 43 (ix) TELECOMMUNICATION INFORMATION:  
 44 (A) TELEPHONE: 415-855-0555  
 45 (B) TELEFAX: 415-845-4166  
 46 (C) TELEX:  
 49 (2) INFORMATION FOR SEQ ID NO: 1:  
 51 (i) SEQUENCE CHARACTERISTICS:  
 52 (A) LENGTH: 260 amino acids  
 53 (B) TYPE: amino acid  
 54 (C) STRANDEDNESS: single  
 55 (D) TOPOLOGY: linear  
 W--> 57 (ii) MOLECULE TYPE: None

## RAW SEQUENCE LISTING

DATE: 08/29/2002

PATENT APPLICATION: US/09/745,506

TIME: 15:33:15

Input Set : N:\Crif3\RULE60\09745506.raw

Output Set: N:\CRF4\08292002\I745506.raw

58 (vii) IMMEDIATE SOURCE:

59 (A) LIBRARY: U937NOT01

60 (B) CLONE: 187

62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

64 Met Pro Gln Asn Glu Tyr Ile Glu Leu His Arg Lys Arg Tyr Gly Tyr

65 1 5 10 15

66 Arg Leu Asp Tyr His Glu Lys Lys Arg Lys Lys Glu Ser Arg Glu Ala

67 20 25 30

68 His Asp Arg Ser Lys Lys Ala Lys Lys Met Ile Gly Leu Lys Ala Lys

69 35 40 45

70 Leu Tyr His Lys Gln Arg His Ala Glu Lys Ile Gln Met Lys Lys Thr

71 50 55 60

72 Ile Lys Met His Glu Lys Arg Asn Thr Lys Gln Lys Asn Asp Glu Lys

73 65 70 75 80

74 Thr Pro Gln Gly Ala Val Pro Ala Tyr Leu Leu Asp Arg Glu Gly Gln

75 85 90 95

76 Ser Arg Ala Lys Val Leu Ser Asn Met Ile Lys Gln Lys Arg Lys Glu

77 100 105 110

78 Lys Ala Gly Lys Trp Glu Val Pro Leu Pro Lys Val Arg Ala Gln Gly

79 115 120 125

80 Glu Thr Glu Val Leu Lys Val Ile Arg Thr Gly Lys Arg Lys Lys Lys

81 130 135 140

82 Ala Trp Lys Arg Met Val Thr Lys Val Cys Phe Val Gly Asp Gly Phe

83 145 150 155 160

84 Thr Arg Lys Pro Pro Lys Tyr Glu Arg Phe Ile Arg Pro Met Gly Leu

85 165 170 175

86 Arg Phe Lys Lys Ala His Val Thr His Pro Glu Leu Lys Ala Thr Phe

87 180 185 190

88 Cys Leu Pro Ile Leu Gly Val Lys Lys Asn Pro Ser Ser Pro Leu Tyr

89 195 200 205

90 Thr Thr Leu Gly Val Ile Thr Lys Gly Thr Val Ile Glu Val Asn Val

91 210 215 220

92 Ser Glu Leu Gly Leu Val Thr Gln Gly Gly Lys Val Ile Trp Gly Lys

93 225 230 235 240

94 Tyr Ala Gln Val Thr Asn Asn Pro Glu Asn Asp Gly Cys Ile Asn Ala

95 245 250 255

96 Val Leu Leu Val

97 260

99 (2) INFORMATION FOR SEQ ID NO: 2:

101 (i) SEQUENCE CHARACTERISTICS:

102 (A) LENGTH: 153 amino acids

103 (B) TYPE: amino acid

104 (C) STRANDEDNESS: single

105 (D) TOPOLOGY: linear

W--&gt; 107 (ii) MOLECULE TYPE: None

108 (vii) IMMEDIATE SOURCE:

109 (A) LIBRARY: U937NOT01

110 (B) CLONE: 2335

112 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

## RAW SEQUENCE LISTING

DATE: 08/29/2002

PATENT APPLICATION: US/09/745,506

TIME: 15:33:15

Input Set : N:\Crf3\RULE60\09745506.raw

Output Set: N:\CRF4\08292002\I745506.raw

```

114 Met Ser Lys Asn Thr Val Ser Ser Ala Arg Phe Arg Lys Val Asp Val
115 1 5 10 15
116 Asp Glu Tyr Asp Glu Asn Lys Phe Val Glu Arg Arg Arg Trp Gly Arg
117 20 25 30
118 Arg Ala Arg Ala Arg Ala Pro Ser Glu Gly Glu Val Asp Ser Cys Leu
119 35 40 45
120 Arg Gln Gly Asn Met Thr Ala Ala Leu Gln Ala Ala Leu Lys Asn Pro
121 50 55 60
122 Pro Ile Asn Thr Lys Ser Gln Ala Val Lys Asp Arg Ala Gly Ser Ile
123 65 70 75 80
124 Val Leu Lys Val Leu Ile Ser Phe Lys Ala Asn Asp Ile Glu Lys Ala
125 85 90 95
126 Val Gln Ser Leu Asp Lys Asn Gly Val Asp Leu Leu Met Lys Tyr Ile
127 100 105 110
128 Tyr Lys Gly Phe Glu Ser Pro Ser Asp Asn Ser Ser Ala Met Leu Leu
129 115 120 125
130 Gln Trp His Glu Lys Ala Leu Ala Ala Gly Gly Val Gly Ser Ile Val
131 130 135 140
132 Arg Val Leu Thr Ala Arg Lys Thr Val
133 145 150

```

135 (2) INFORMATION FOR SEQ ID NO: 3:

137 (i) SEQUENCE CHARACTERISTICS:

138 (A) LENGTH: 185 amino acids

139 (B) TYPE: amino acid

140 (C) STRANDEDNESS: single

141 (D) TOPOLOGY: linear

W--&gt; 143 (ii) MOLECULE TYPE: None

144 (vii) IMMEDIATE SOURCE:

145 (A) LIBRARY: HUVENOB01

146 (B) CLONE: 36079

148 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

150 Met Phe Ala Asp Thr Leu Leu Ile Val Phe Ile Ser Val Cys Thr Ala
151 1 5 10 15
152 Leu Leu Ala Glu Gly Ile Thr Trp Val Leu Val Tyr Arg Thr Asp Lys
153 20 25 30
154 Tyr Lys Arg Leu Lys Ala Glu Val Glu Lys Gln Ser Lys Lys Leu Glu
155 35 40 45
156 Lys Lys Lys Glu Thr Ile Thr Glu Ser Ala Gly Arg Gln Gln Lys Lys
157 50 55 60
158 Lys Ile Glu Arg Gln Glu Glu Lys Leu Lys Asn Asn Asn Arg Asp Leu
159 65 70 75 80
160 Ser Met Val Arg Met Lys Ser Met Phe Ala Ile Gly Phe Cys Phe Thr
161 85 90 95
162 Ala Leu Met Gly Met Phe Asn Ser Ile Phe Asp Gly Arg Val Val Ala
163 100 105 110
164 Lys Leu Pro Phe Thr Pro Leu Ser Tyr Ile Gln Gly Leu Ser His Arg
165 115 120 125
166 Asn Leu Leu Gly Asp Asp Thr Thr Asp Cys Ser Phe Ile Phe Leu Tyr
167 130 135 140

```

## RAW SEQUENCE LISTING

DATE: 08/29/2002

PATENT APPLICATION: US/09/745,506

TIME: 15:33:15

Input Set : N:\Crif3\RULE60\09745506.raw

Output Set: N:\CRF4\08292002\I745506.raw

```

168 Ile Leu Cys Thr Met Ser Ile Arg Gln Asn Ile Gln Lys Ile Leu Gly
169 145                               150                155                160
170 Leu Ala Pro Ser Arg Ala Ala Thr Lys Gln Ala Gly Gly Phe Leu Gly
171                               165                170                175
W--> 172 Pro Pro Pro Xaa Ser Gly Lys Phe Ser
173                               180                185
175 (2) INFORMATION FOR SEQ ID NO: 4:
177 (i) SEQUENCE CHARACTERISTICS:
178 (A) LENGTH: 106 amino acids
179 (B) TYPE: amino acid
180 (C) STRANDEDNESS: single
181 (D) TOPOLOGY: linear
W--> 183 (ii) MOLECULE TYPE: None
184 (vii) IMMEDIATE SOURCE:
185 (A) LIBRARY: HUVESTB01
186 (B) CLONE: 82709
188 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
190 Met Ser Thr Asn Asn Met Ser Asp Pro Arg Arg Pro Asn Lys Val Leu
191 1 5 10 15
192 Arg Tyr Lys Pro Pro Pro Ser Glu Cys Asn Pro Ala Leu Asp Asp Pro
193 20 25 30
194 Thr Pro Asp Tyr Met Asn Leu Leu Gly Met Ile Phe Ser Met Cys Gly
195 35 40 45
196 Leu Met Leu Lys Leu Lys Trp Cys Ala Trp Val Ala Val Tyr Cys Ser
197 50 55 60
198 Phe Ile Ser Phe Ala Asn Ser Arg Ser Ser Glu Asp Thr Lys Gln Met
199 65 70 75 80
200 Met Ser Ser Phe Met Leu Ser Ile Ser Ala Val Val Met Ser Tyr Leu
201 85 90 95
202 Gln Asn Pro Gln Pro Met Thr Pro Pro Trp
203 100 105
205 (2) INFORMATION FOR SEQ ID NO: 5:
207 (i) SEQUENCE CHARACTERISTICS:
208 (A) LENGTH: 166 amino acids
209 (B) TYPE: amino acid
210 (C) STRANDEDNESS: single
211 (D) TOPOLOGY: linear
W--> 213 (ii) MOLECULE TYPE: None
214 (vii) IMMEDIATE SOURCE:
215 (A) LIBRARY: LUNGNOT02
216 (B) CLONE: 313727
218 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
220 Met Ala Ala Ile Pro Pro Asp Ser Trp Gln Pro Pro Asn Val Tyr Leu
221 1 5 10 15
222 Glu Thr Ser Met Gly Ile Ile Val Leu Glu Leu Tyr Trp Lys His Ala
223 20 25 30
224 Pro Lys Thr Cys Lys Asn Phe Ala Glu Leu Ala Arg Arg Gly Tyr Tyr
225 35 40 45
226 Asn Gly Thr Lys Phe His Arg Ile Ile Lys Asp Phe Met Ile Gln Gly

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/745,506

DATE: 08/29/2002

TIME: 15:33:15

Input Set : N:\Crif3\RULE60\09745506.raw

Output Set: N:\CRF4\08292002\I745506.raw

```

227      50      55      60
228 Gly Asp Pro Thr Gly Thr Gly Arg Gly Gly Ala Ser Ile Tyr Gly Lys
229 65      70      75      80
230 Gln Phe Glu Asp Glu Leu His Pro Asp Leu Lys Phe Thr Gly Ala Gly
231      85      90      95
232 Ile Leu Ala Met Ala Asn Ala Gly Pro Asp Thr Asn Gly Ser Gln Phe
233      100     105     110
234 Phe Val Thr Leu Ala Pro Thr Gln Trp Leu Asp Gly Lys His Thr Ile
235      115     120     125
236 Phe Gly Arg Val Cys Gln Gly Ile Gly Met Val Asn Arg Val Gly Met
237      130     135     140
238 Val Glu Thr Asn Ser Gln Asp Arg Pro Val Asp Asp Val Lys Ile Ile
239      145     150     155     160
240 Lys Ala Tyr Pro Ser Gly
241      165
243 (2) INFORMATION FOR SEQ ID NO: 6:
245 (i) SEQUENCE CHARACTERISTICS:
246 (A) LENGTH: 173 amino acids
247 (B) TYPE: amino acid
248 (C) STRANDEDNESS: single
249 (D) TOPOLOGY: linear
W--> 251 (ii) MOLECULE TYPE: None
252 (vii) IMMEDIATE SOURCE:
253 (A) LIBRARY: BRSTNOT05
254 (B) CLONE: 965366
256 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
258 Met Ala Lys Gln Gly Gln Met Asp Ala Val Arg Ile Met Ala Lys Asp
259 1      5      10      15
260 Leu Val Arg Thr Arg Arg Tyr Val Arg Lys Phe Val Leu Met Arg Ala
261      20      25      30
W--> 262 Asn Ile Gln Ala Val Phe Leu Lys Ile Gln Thr Xaa Lys Phe Asn Xaa
263      35      40      45
W--> 264 Xaa Met Ala Gln Ala Met Lys Gly Val Asn Lys Ala Met Gly Thr Met
265      50      55      60
266 Asn Arg Gln Leu Lys Leu Pro Gln Ile Gln Lys Ile Met Met Glu Phe
267 65      70      75      80
W--> 268 Glu Arg Gln Ala Glu Ile Met Xaa Met Lys Glu Glu Met Met Asn Asp
269      85      90      95
270 Ala Ile Asp Asp Ala Met Gly Asp Glu Glu Asp Glu Glu Glu Ser Asp
271      100     105     110
272 Ala Leu Val Phe Gln Gly Leu Asp Glu Leu Gly Leu Ser Leu Thr Asp
273      115     120     125
274 Glu Leu Ser Asn Leu Pro Ser Thr Gly Gly Ser Leu Ser Val Ala Ala
275      130     135     140
276 Gly Gly Lys Lys Ala Glu Ala Ala Ala Ser Ala Leu Ala Asp Ala Asp
277      145     150     155     160
278 Ala Asp Leu Glu Glu Arg Leu Lys Asn Leu Arg Arg Asp
279      165     170
281 (2) INFORMATION FOR SEQ ID NO: 7:

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/09/745,506

DATE: 08/29/2002  
TIME: 15:33:16

Input Set : N:\Crf3\RULE60\09745506.raw  
Output Set: N:\CRF4\08292002\I745506.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; Xaa Pos.180  
Seq#:6; Xaa Pos.44,48,49,88  
Seq#:7; Xaa Pos.210,212,224,230  
Seq#:9; Xaa Pos.175  
Seq#:12; Xaa Pos.85  
Seq#:13; Xaa Pos.194  
Seq#:14; Xaa Pos.45  
Seq#:16; Xaa Pos.9  
Seq#:21; Xaa Pos.3,15  
Seq#:38; N Pos. 1114  
Seq#:39; N Pos. 1266,1327  
Seq#:43; N Pos. 263,265,277,279,396,729  
Seq#:44; N Pos. 755,762,796,816,858,897,906,910  
Seq#:45; N Pos. 8,13,20,36,41,95,98,101,104,106  
Seq#:46; N Pos. 17,24,620,820  
Seq#:47; N Pos. 839,840,841  
Seq#:49; N Pos. 18,333,653  
Seq#:50; N Pos. 607  
Seq#:51; N Pos. 272  
Seq#:52; N Pos. 583  
Seq#:53; N Pos. 37  
Seq#:54; N Pos. 517,540,581,586,599,615,626,627,628,629,632,635,642,643,647  
Seq#:54; N Pos. 656,664,665,671,681,687,692,693,695,706,710,713,715,716  
Seq#:56; N Pos. 44,47,51,659  
Seq#:58; N Pos. 55,90  
Seq#:64; N Pos. 64  
Seq#:71; N Pos. 613

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/745,506

DATE: 08/29/2002

TIME: 15:33:16

Input Set : N:\Crf3\RULE60\09745506.raw

Output Set: N:\CRF4\08292002\I745506.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]  
L:57 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=1  
L:107 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=2  
L:143 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3  
L:172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:176  
L:183 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=4  
L:213 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5  
L:251 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6  
L:262 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:32  
L:264 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:48  
L:268 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:80  
L:289 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7  
L:322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:208  
L:324 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:224  
L:337 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8  
L:379 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9  
L:406 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9 after pos.:160  
L:423 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10  
L:463 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=11  
L:493 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=12  
L:510 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12 after pos.:80  
L:531 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=13  
L:562 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:192  
L:579 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=14  
L:590 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14 after pos.:32  
L:629 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15  
L:667 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=16  
L:674 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:0  
L:695 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=17  
L:723 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=18  
L:773 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=19  
L:807 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=20  
L:845 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=21  
L:852 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21 after pos.:0  
L:873 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=22  
L:907 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=23  
L:937 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=24  
L:971 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=25  
L:1001 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=26  
L:1057 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=27  
L:1091 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=28  
L:1123 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=29  
L:1157 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=30  
L:1213 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=31  
L:1257 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=32  
L:1287 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=33  
L:1337 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=34

## VERIFICATION SUMMARY

DATE: 08/29/2002

PATENT APPLICATION: US/09/745,506

TIME: 15:33:16

Input Set : N:\Crf3\RULE60\09745506.raw

Output Set: N:\CRF4\08292002\I745506.raw

L:1373 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=35  
L:1409 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=36  
L:1443 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=37